

Sequence Information.

Sequence ID No. 1: ATR seq

1 GCGCTCTTCCGGCAGCGGTACGTTGGAGACCCGGGAACCCGGTTGGCGTGGTGAATGGCTCGCAGCCT 75
 5 76 CAGCATGGGGAAACATGGCTGGAGCTGGCTTCATGATCCCCGCCCTGGGGAGCTGGCAGTGCACACCAGA 150
 151 GGAATATAATACAGTTGACAGAACGCAAGACAATTCTGTCAATTCTGACCCGAAACTACAGATGAAA 225
 226 TGTGTTGCTGAGAACCTGACTCTCACGCCAACCTCGTGATGTTGCTTGTATTCCAGCA 300
 301 TATCATGAAATCTCCCACTTATGTTGAAATGTGAGTGGAGCCATGAGCGCAAAGGCAGTTGATTGAATT 375
 376 CAGTAATTGGATCATAACGAGACTCTGGGATTGAGCAACTCCCTCTGTCAATTGTTACACAAGAAAATCTG 450
 451 TGAAGTCATCTGTCATTATTATTCTTTAAAAGCAAGAGTCTGCTATTTGGGTACTCACAAAAGAATT 525
 10 526 ATTACAACCTTTGAAAGACTTGGTTACCTCATAGAAGAAATGTGATGGGTATGCTGTGGAATGGCCAGTGGT 600
 601 CATGAGCCGATTTTAAGTCATTAGATGAACACATGGGATATTACATCAGCTCCTTGAGTTGATGAGTAT 675
 676 GCAAAATTAGAATTATTGAAGTCACCTTATTAAATGGTTTACTCGTATTATTGCAATTGTTGTTTAAAG 750
 751 GCAAGAACCTTACTTGGCAGATAGGTTGTTCTGCTAGAGTATGGTAGTCCAAAATTAATCCCTAGCAAT 825
 826 TAGTTTTAACAGAACCTTCACTGGAGGACTACCCAGCACACCCAGCTAGCACTTTTCAGCTCATTTT 900
 15 901 GGAATTATTAAAACACCTTGAGAACATGGGATACTGACCAATTGAAACTCTATGAAGAGCATTATCAAAGCTGAT 975
 976 AAAGACACTATTCCTTGAAGCAGAACGTTAGAAATTTGAACTGCTATTAAATATGCTGCTGGAAAA 1050
 1051 ACTCTGTGTCATTTGAAAGACGGTGTGCTCATGGCTTAAGTCTGATTGCTAAAGCAGCTTGCCATT 1125
 1126 ACTGCGATTTCTTAAATTGTGCCAGCTGGTATGAACCTGCTTACAAGTCAGGAAGGTCTATGTGAGAAA 1200
 1201 TATTGAAAGCTTTGGATGTGCTTGGAAATTGAGGTAGATGCAAGAGTACTTGTGGGCCACTTATGCAGC 1275
 20 1276 TTGAAAATGGAAAGTATGGAAATCATTGGAGGATTCAATGCCAAACTCAACAGGAAACCTCAGCAGTAATAG 1350
 1351 TGATGGAATATCACCCAAAAGCGTCGCTCAGCTGCTCTAAACCTCTAAAGAGCACCACAGACTGA 1425
 1426 GGAAATTAAACATGTGGACATGAACCAAAAGAGCATATTATGGAGTGCAGACTGAAACAGAAAGCTGAATCCCTCA 1500
 1501 GATTCCCTGAATACAGTGGCTAAAGAACCTGTATTGAGATGTTAGAAGGATTGCTGTGCTTACAAC 1575
 1576 GACTGCTCTGTACTGTCATTGTCATCAAAACATGAACCTGGTACTTCAAGGACTGTCACATCAAATC 1650
 25 1651 CAAGAAGAAACCTCTGTAGTGATAACCTGGATGTCATTGGATTTCACACAAAGTCTTAAGAGCTGTAGAAG 1725
 1726 TTGTTAGAATCTGTCAGAACACTGGACCTGGAGGCAACATTGATAAGGTGGTAAAATTATGATGCTTGT 1800
 1801 TTATATGCAAGTAAACAGTCATTGAAAGATCATATCCTGGAAGATTATGTTGCTTACTCCATGGAT 1875
 1876 TTATCCCATTCTGATGGCTGTTAAAGTTGACCACATTGGCGCTAATCTCAACATTAAAGCTGTAGGGT 1950
 1951 TTCAGATAGTATTCAACCACAGGCACAACTCACGATGTTGTTCTCTGACTCTGTTCCAAGAAGAATTCT 2025
 30 2026 TGAGTGGAGAACAGCAGTTACAACCTGGGCCCTGAGAGCTCCCATGAAGTAATCGGGCTAGTTGTTAGTGG 2100
 2101 ATTTTATCTTATTGAGCAGCAGAACATTCTGTAACAGAGTCCCAAGATTCTTATAGATAAAAGTCAGATGA 2175
 2176 TTCTGACATTGTCAAGAACGAAATTGCTTCTACTTGGTCAACCTGCTGACTCTCACGGCATGTTTATCT 2250
 2251 GACAAGTTTAAACAGAACCTCTCTGAACACGGACATGGACCTCTCTGTAGGAACCTGAAAGCCACTTC 2325
 2326 TCAACATGAATGTCATCTCTCAACTAAAGCTCTGTCTGAAGCCATTCTTCTACTGAAAAAAAT 2400
 35 2401 ACCTAGTCCAGTAAACACTGCTTCAAGATAATCTACATCATCTTGTAAAGCATCTGATTAGAGATGA 2475
 2476 AACAGATGTAAGAACGAGTTCTGGAACTTAAATTAAATGAAAGATCCAGACAAAGATGTTAGAGTGGCTT 2550
 2551 TAGTGGAAATATCAAGCACATATTGAAATCCTGGACTCTGAAGATGGATTATAAAGGAGCTTTGTCTTAAAG 2625
 2626 AATGAAGGAAGCATATACACATGCCAAATATCAAGAACATAATGAGCTGAAGGATACCTGATTCTACAACAGG 2700
 2701 GGATATTGGAAGGGCGCAAAAGGAGATTGGTACCATTTGCACTCTTACACTTATTGCAATTGTTATCCAA 2775
 40 2776 GTCAGCATCTGCTCTGGAGCAGCATAACAGAACATTAGAGCTCTGGTGCAGCTAAAGTGTAAACTGCAAAG 2850
 2851 TTTTCTAGCCAGTATAAGAACCCAATGTCAGTTTGGTAGAATCCCTCACTCTAGTCAGATGACAGCACT 2925
 2926 TCCGAATACTCCATGCCAGAACGTCAGTGGCAAAACAAGATGTTGCTCACCAGAGAGAAATGGCTTAAAC 3000
 3001 GTTGTCTGAAATTGCCAACGTTTGCACCTTCTGATCTTAACTGTTTCTACTAGGACATTACAAGTTACT 3075

- 40 -

3076 ACCTGATCTTGTGCCAAGCAAGCCCTGCAGCTCTGCTCTCATTGAACCTTAGGAAAACAATTAAATGTCAA 3150
 3151 TCGTAGAGAGATTTAATAAACAACTTCAAAATATATTTCTCATTTGGTCTGTTCTGTTCCAAGATGAATT 3225
 3226 AGAACGTGCCCTCATTATCTGAAGAATGAAACAGAAAATGAACTGGGAGCCTGTTGAGACAAGATTCCAAGG 3300
 3301 ATTGCATAATGAATTATTGCTCGTATTGAGAACACTATCACAGGTTTAATGGTTGTCAATACTGCCCTC 3375
 5 3376 ATTGCGATCCAGTGTGATCCATATCAGGGCCCAGAGAGATATCATATCACCTGAACGTGATGGCTGATTATTACA 3450
 3451 ACCCAAATTGTTGGCATTGGCTTTTAACATGCGTTACTGAGCTCTAGTGTGTCAGGTTGTGAGGGTGAAGAT 3525
 3526 AATGGCCTTGAAAGCTTGTCTTGTGATGAAGTTAATGGGACCCAAACATGTCAGGTTGTGAGGGTGAAGAT 3600
 3601 GATGACCACACTGAGAACACTGGCCTTCGATTAAGGATGATTTCCTGAATTGTGTCAGGCTGGGACTGCTT 3675
 3676 TGTCGCTGCCCTGGATCATGCTTGTCTGGCTCCCTCTCAGTCATGTAATAGTAGCTTGTACCTTATACA 3750
 10 3751 CATCCAGCTAAAGAAACTGCAGCTATCTTCACTACCTCATAATTGAAAACAGGGATGCTGTGCAAGATTCT 3825
 3826 TCATGAAATATATTTTACCTGATCATCCAGAATTAAAAAGATAAAAGCCGTTCTCCAGGAATACAGAAAGGA 3900
 3901 GACCTCTGAGAGACTGATCTCAGACAACCTCTCAGCTCTCATGAAGGCCATTCAACATGAAAATGTCGATGT 3975
 3976 TCGTATTCTCATGCTTACAAGCTGAAGGAAACCTTGTATAAAATCAGGAAAACGATAAAAGTATGCAACAGA 4050
 15 4051 CAGTGAACAGTAGAACCTTATATCTCACAGTTGGTACAGTGTCTTGAAGGTTGCAAGATGCAAACCTCTCA 4125
 4126 AGCTCGGTTGCTGTGGGAATGTTAGGGATTGGGGCAGATAGATCCAGGTGATTAGATTCTCAACAAAC 4200
 4201 TGAAACTCAAGGAAAAGATTACATTGTGACTGGAGTAGAAGGATTCAGCTTGTGATTGATTTGATGGA 4275
 4276 GCTAACAAAGAGCTTACCTTGCATGCTGATAATAGCCGAGCTCAAGATTGCTGCCTATGCCATTAGGAGTT 4350
 4351 GCTTCTTATTTATGACTGTAGAGAGATGGAGACCAACGGCCAGGTACCCAATTGTGGAGGAGATTCTGAGCA 4425
 4426 TGTCGGGAAATACTAGAACCTCATCTAAATACCAAGATAACAGATTCTCAGAAGTCACCGATTGGTCTGGAGT 4500
 20 4501 AAAGAAGCCAATTACTTAAGTAAATTGGTAGTAACTTGCAAGATGGTCAGCATCTGGCAGGTTATCTTAT 4575
 4576 TACAAAGGTTGCACATGATCTGCCAGTAAAATTTCACCTGCTGAGCATTATGATGAAGCATGATTCAAAGT 4650
 4651 GACCATCTATCTTCCACATATTCTGGTGTATGCTTACTGGGTTGTAATCAAGAAGATCAGCAGGAGGTTA 4725
 4726 TGCAAGAAATTATGGCAGTTCTAAAGCATGACGATCAGCATACATAAAACCAAGACATTGATCTGATCTGTG 4800
 4801 TCAACTCAGTACACAGACTGTGTTCTCATGCTTGACCATCTCACACAGTGGCAAGGCACAAATTCAAGGCACT 4875
 4876 GAAAGCTGAGAAATGTCACACAGCAAATCAAACAGAAATAAGGTAGACTCAATGGTACTGTGATTATGA 4950
 4951 AGACTATCAGAGTGTAAACCGTTCTAGACCTCATACCCAGGAACTCTGGCAGTAGCTTCTCCTGCTCCAA 5025
 5026 AGCATACACACGAGCTGTAATGCACTTGAATCATTTATTACAGAAAAGCAAATTCAGGAACATCTTGG 5100
 5101 ATTTTACAGAAAATTGTATGCTGATGCAACCTGATGGAGTGGCCGGAGTCAGTGCATTAGAAAGGCAGA 5175
 5176 ACCATCTCTAAAGAACAGATCTTGAACATGAAAGCCTGGCTTGCTGAGGGATGCCACTGCTTGTGACAG 5250
 30 5251 GGCTATTCAAGCTAGAACCAAGACAGATCATTATCATGGTAGTAAAGTCCATGTTAGGTCTGGTCAGCT 5325
 5326 GTCTACTGTTATCACTCAGGTGAATGGAGTGCATGCTAACAGGTCCGAGTGGACAGATGAATTAAACACGTACAG 5400
 5401 AGTGGAAAGCAGCTTGGAAATTGTCACAGTGGGATTGGTGGAAAACATTGGCAGCAGATGGAAAATCTACAAC 5475
 5476 ATGGAGTGTGAGACTGGGACAGCTTATTATCAGCCTTAAAGAGATATCACAGCTTATGACTCACTGAA 5550
 5551 ACTAGTGGAGGAGAACAAATTGACCTCTTCAAGCTGCAAGCTTGAAGAGGCTCCTACCAACGAGGAGATGAA 5625
 5626 ATATATTGTGAGATTGACATGTTATGTGAGTTGGAGCATAGCATCAAACACTTTCCAGCATCTCAGGTGA 5700
 5701 CAGTTCTCAAGAACAGATTCTAACTGGTAGCTGCACTAGAAAATGACCCAGAATTCTACAGAGCCAAGGAGCC 5775
 5776 TATCCTGGCTCTCCGGAGGGCTTACTAAGCTCAACAAAAGACAGATTACAATGAAATGGTTGGAGAATGCTG 5850
 5851 GCTGCAGAGTGCAGGGTAGCTAGAAAGGCTGGTACCCACAGACAGCCTACAATGCTCTCTTAATGCAAGGGGA 5925
 5926 ATCACCACTCGCTGAACGTACGGAAAGGGCAAAGTGGCTGGTCAAGGGTGTGTTACCAAGGCACTAAT 6000
 40 6001 TGTTCTCAAAAAGGTGTTGAATTATGTTTCTGAAAATGAAACCCACCTGAGGGTAAGAACATGTTAATCCA 6075
 6076 TGGTCGAGCTATGCTACTAGTGGGCCATTATGGAAGAAACAGCTAACATTGAAAGCAATTGAAAAAA 6150
 6151 ATATAAGGATGTGACCGCGTGCCTGCCAGAATGGGAGGATGGCATTTCACCTGCAAGTACTATGACAAATT 6225
 6226 GATGCCCATGGTCACAGACAACAAATGGAAAAGCAAGGTGATCTCATCCGTATATAGTTCTCATTTGGCAG 6300
 6301 ATCTCTACAATATGAAATCAGTCATATATCAGTCATGCCACGAATGTTAACTCTATGGCTTGTGATTATGGTAC 6375
 45 6376 AAAGGCATATGAATGGAAAAGCTGGCCGTCGATCGTACAAATGAGGAATGATTGGTAAAATAACAA 6450
 6451 GTGTCATTCTCACGATGAAAGTTTGTGTTGATGAAATAAGCCAAGTATTCTAGCCTATCCTCAACA 6525
 6526 TTGTCATTCTCACGATGAAAGTTTGTGTTGATGAAATAAGCCAAGTATTCTAGCCTATCCTCAACA 6600

- 41 -

6601 AGCAATGTGGATGATGACAGCTGTCAAAGTCATCTTATCCCATCGCTGTGAACAGATGCAAGGAATCCTCAA 6675
 6676 TAAAGCTATTATGAAAAATCCTAGAGAAGTTGGAGATGCAACTCGCTAACAGATAAGCTCTAGA 6750
 6751 ATTGTGCAATAAACCGGTTGATGGAAGTAGTCCACATTAAGCATGAGCACTCATTTAAAATGCTAAAAAGCT 6825
 6826 GTAGAAGAACACATTAGTGAATCCTCATTCTACATCAGTCATGACCTACACTCCATCAATTCT 6900
 5 6901 GGGTACCCATGCTAACCATGCTAGGCATGAACCATTCCTGGACATGGCTATATTGAGGGTTGATGATAT 6975
 6976 GGTGAAATTCTGCTCTCAGAAACCAAAGAAGATTCTTAAAGGCTAGATGAAAGTTACATCAT 7050
 7051 GATGTGAAAGCAGGTTGAGACTAATGGAATTCAATTCTGATTAATAAGTGCTT 7125
 7126 AAGAAAAGATGCAGAGTCTCGTAGAAGAGAACCTCATATTGAAACATATGAGCTTATTCCACTAAATGATGAATG 7200
 7201 TGGGATTATTGAATGGGTGAAACAACACTGCTGGTTGAGACCTATTCTGACCAAACATATAAAGAAAAGGGAGT 7275
 10 7276 GTATATGACAGGAAAAGAACCTTCGCCAGTGTATGCTACCAAAGTCAGCACCTTATCTGAAAAACTCAAAGTATT 7350
 7351 CCGAGAAATTCTCTGCCAGGCATCCTCTATTTCATGAGTGGTTCTGAGAACATTCCCTGATCCTACATC 7425
 7426 ATGGTACAGTAGATCAGCTTACTGCCGTTCACTGCACTGAGTAATGCAATGGTTGGTTATATTCTGGGGCTTG 7500
 7501 AGACCGTCATGGTAAAGATATTCTCTTGTACTGGTGAATGGTACATGTAGATTCAATTGTCTTT 7575
 7576 CAATAAGGGAGAACACCTTGAAGTCCAGAAATTGCGCATTCGCTGACTCATAATATGGTTAATGGAATGGG 7650
 15 7651 TCCTATGGGAAACAGAGGGCTTTGCAAGAGCATGTGAAGTTACAATGAGGCTGATGCGTGTACAGCAGAGCC 7725
 7726 TTAAATGAGTGTCTAAAGACTTTCTACATGATCCTCTGTGGAATGGAGTAACAGCTGAAAGGGCATTCCAA 7800
 7801 AGGCCACTGAATGAAACTGGAGAAGTTGCAATGAAAGGCCAACGACCATGTTCTGACATTGAGCAGCGACT 7875
 7876 ACAAGGTGAATCAAGACTGAAATAGAGTGCAGGGACTGCCGTTATGAAAGGACATGTGCAATTACCTTAT 7950
 7951 ACAAGAAGCTACTGATGAAACTACTATGCCAGATGATCTGGTTGGACTCCATATAATGAAATGAAATTAT 8025
 20 8026 GTAAAAGAATATGTTAATAATCTAAAGTAATGCATTGGTATGAACTGTGGTTGTATGTTCAATTCTAAAG 8100
 8101 TACAACATAAAATTACGTTCTCAGCACTGTTATTCTCTGATCATTAATTATGTTAAATATACATT 8175
 8176 AGTTATTAAGAAATAACTGCTTCTTAATAAAAAAAAAAAAAAAAAAAAAAA 8239

Sequence ID No. 2: ATR protein

1 MGEHGLEASMI PALRELGSATPEEYNTVVKPROILCQFIDRILTDVN VVAVELVKKTDSOPTSV 66
 25 67 MLLDFIOHIMKSSPLMFNVSGSHERKGSCIEFSNWIITRLRIAATPSCHLHKKICEVICSLLFLFKSKSPAI 141
 142 FGVLTKELLQLFEDLVYLHRRNVMGHAVEWPVMSRFLSQLDEHMGLQSAPLQLMSMQNLEFIETLMLMVLTRI 216
 217 IAIFFRROEPLLWQIGCVLLEYGSPKIKSLAISFLTEFOLGGPAQPASTFFSSFLELLKHLEMVEMTDQKL 291
 292 EEPPLSLKLTKLFPFEAEAYRNIEPVYNMNLLEKLCVMFEDGVLMRLKSDLKAALCHLLOYFLKFVPAGYESALO 366
 367 VRKVYVRNICKALLDVLGIEVDAEYLLGPLYAALKMESMEIEEIOCOTQENLSSNSDGISPKRRLSSSLNPS 441
 442 KRAPKOTEEIKHVDMNQKSILWSALKQKAESLQISLEYSGLKNPVIENLEGIAVVLOLTALCTVHC SHONMCRT 516
 517 FKDCOHKSKKPKS VVITWMSLDFYTKVLKCSRSLLESVQKLDLEATIDKVVKIYDALIYMQVNSSFEDHILEDLC 591
 592 GMLSLPWIYSHSDGCLKLTTFAANLLTLSRISDSYSPQAQSRCVFLTLFPRRIFLEWR TAVYNWALOSSHEV 666
 667 IRASCVGFFILLQQQNCSNRVPKILJDVKDDSDIVKEFASILGOLVCTLHGMFYLTSSLTEPFSEHGHVDLF 741
 742 CRNLKATSQHECSSSQLKASVCKPFLFLLKKKIPSPVKA FIDNLHHLKHDREDT DVKA LGTLLNLMDP 816
 35 817 DKDVRVAFSGNIKHL ELSLDSEDFIKEFVLRMKEAYTHAOISRNNEKD TLTTGDIGRAKGDLVPFALLH 891
 892 LLHCLLSKSASVSGAAYTEIRALVAAKSVKLOSFFSOYKKPQCFLVESLHSQMTALPNTPCONADVRKQDV AH 966
 967 OREMALNTLSEIANVFD PDLNRF LTRTLOVLPDLAAKASPAASALIRTGKQLNVNRREILINNFKYIFSHLV 1041
 1042 CSCSKDELERALHYLKNETIELGSLLRQDFQGLHNELLRLIGEHYOOVFNGLSILASRASSDDPYQGPDIISP 1116
 1117 ELMADYLQPKLLGILA FFNMOLLSSSSVGIEDK KMA NLSLMSLMKLMGPKHVSSVRVKMMT LRTGLRFKDDFPEL 1191
 40 1192 CCRAWDCFVRCLDHACLGSLLSHVIVALLPLIHIOPKETAAIFHYLI IENRDAVQOF LHIEIYFLPDHPELKKIKA 1266
 1267 VLQEYRKETSESTDLOLTTQLSMKAIOHENVDVRIHALTSLKETLYKNQEKLIK YATDSETVEPIISOLVTVLK 1341
 1342 GCQDANSQARLLCGECLGELGAI DPGRLDFSTTETOGKDTFVTGVEDSSFAYGLLMELTRAYADNSRAQDS 1416
 1417 AAYAIQELLSIYDCREM:TNGPGHOLWRRFPEHVREILEPHLNTRYKSSOKSTDWSGVKKP IYLSKLGNSFAEWS 1491
 1492 ASWAGYLITKVRHDLASKIFTCCS IMM KHDFKVI:YLLPHILVYVLLGCNOEDQEVYAEJMAVLKHDDOHTINT 1566
 45 1567 QDIASDLCOLSTOTVFSMLDH_TOWARHKFOALKAEKCPHSKSNRNKVDSMVSTVDYEDYOSVTRFLDL IPODTL 1641

- 42 -

1642 AVASFRSKAYTRAVMHFESFITEKKQNIQEHLGFLQKLYAAMHEPDGVAGVSAIRKAEPSLKEQILEHESLGLLR 1716
 1717 DATAACYDRAIQLEPDQIHYHGVVKSMLGLGQLSTVITQVNGVHANRSEWDELNTYRVEAAWLSQWOLVENYL 1791
 1792 AADGKSTTWSVRLQGQLLLSACKRDITAFYDSLKLVRDAOIVPLSAASFERGSYORGYEYIVRLHMLCELEHSIKP 1866
 1867 LFQHSPGDSSQEDSLNWRARLEMTONSYRAKEPILARALLSLNKRDPDNEMVGECHLQSARVARKAGHQTA 1941
 5 1942 NALLNAGESRLAELYVERAKWLWSKGVDVHOALIVLQKGVELCPENETPPEGKNMLIHGRAMLLVGRFMEETANF 1016
 2017 ESNAIMKKYKDVTACLPWEEDGHFYLAKEYDKLMPMVTDNKMKEQGDLIRYIVLHFGRSLQYGNQFIYQSMRML 2091
 2092 TLWLDYGTAKAYEWKEAGRSDRVQMRNDLGKINKVITEHTNYLAPYQFLTAFSQLISRICHSDEVFVVLMEIIAK 2166
 2167 VFLAYPQQAMMMTAVSKSSYPMRVRNCIEILNKAIHMKSLEKFVGDATRLTDKLELCNKPVDGSSSTLSMST 2241
 2242 HFMLKKLVEEATFSEIILPLQSVMIPTLPSILGTHANHASHEPFPGHWAYIAGFDOMVEILASLQPKKISLKG 2316
 10 2317 SDGKFYIMMCKPKDOLRKDCRMEFNSLINKCLRDAEERRRELHIRTAYAPILNDECGIIEWVNNTAGLRPILT 2391
 2392 KLYKEGVYMTGKELROCMPLPKSAELSEKLKVFRFLPRHPPIFHWFRLRTFPDPTSWYSSRSAYCRSTAVMSM 2466
 2467 VGYILGLGDRHGENILFDSTLGECVHDFNCLFNKGTEFEVPEIVPFLRTHNMVNGMGPHTEGLFRRAVEVTMR 2541
 2542 LMRDQREPLMSVLKTFLHDPLVENSXPVKHSKAPLNETGEVNEAKTHVLDIEQRLOGVIKTRNRVTGLPLSI 2616
 2617 EGHVHYLIQEATDENLLCQMYLGWTPYM 2664

15 Sequence ID No. 3: rad3.seq

1 GGTACCAAGTAAAACGTCTTAGTAAGTATAAACACAGAAGAATCCCGATCTAGTGAAACCAATGCCCTGCGTA 75
 76 TGACGCTCCACTGACGCTATAGTCATGAGAACTAGGATGTGCGATTATAACTTATTTCAATATTTCTTAT 150
 151 TATTTATTTAAGAAATAATTGAATTAAAACTCATTTCTTTATTAGCGTAAAATAGCTTATTTCTCTCCT 225
 226 ACTACCTTCACAATAACTTTTTTTGTTATTGACCATATAATCACATCAAAGTCAAAAATTCAATCA 300
 20 301 TTATCAGAAACATCCAGCTTAATATTACTAAAAGTTAGTTCTCTGAAAATTCACTACAAAAGCTCGTTA 375
 376 ATTAGCATCGCTCGATACTTAGTCGACCATGCATCTCCCTTACCTCTGAGTGGAAATGATTTGATAATCGAT 450
 451 TGCCACTTTGCATAATTCTATTGAGATTTTACTTACAATCGCTTTATAATGCTCAAGACTTTGAA 525
 526 CGCGCGTGTGGCTTTAAAAGGCCCTTTTGAATTGAATCAATGGTTGATATAGTATGAGCCAACACGCAA 600
 601 AAAGGAAAGCTGGGTCACTCGATCTTACCCCAGAGGCTTAGATGACAGACAGGTTGGACAGCTTGTAAAG 675
 676 AAGTATTAGCATTAGACAAGAACATGAGTTAGGTAGAAGTAATTCTTACCATCTATGACCTCCGAGCTTGTG 750
 751 AAGTTTTAATTGAAGTTGGCTCTAGCTTTAAACATGATGATTCAAATCTGAATTATCTCTCTAAGATGC 825
 826 TAAAAGAAGCCCATCTCTCTACAAGCGTTAATGCTAATCTTAAAAGGTCTCCGACAGTTTGGGGAGATTA 900
 901 AATCATCTGTTACTCTTGGATTGGATTTACCCAGGACTATATCATTGTTGCTGATATTGTTTATTAAAGT 975
 976 TATTGACTCATTAAAGAGTTCTAAAGCTAATTTATCAGCTAATCAGTCAAAGTCATTCTATGGGACTTAT 1050
 1051 ATGCTTCGTTATGCGTTATGAAATATTATACAAACGTTCTCTATAGTTCTCAAATCAACTATGCTA 1125
 1126 CATTCCCTTACAAGATGCCCTACCCATTCTCAACCATTGCAAGAGTATCTCCCAAATTATCAACCCATCGAG 1200
 1201 AGGACAAATTGATTTACTTATCATTAAATAGAGGGCTTGTACATTCTTCTGAAAGTGCCATTTTTG 1275
 1276 CACAATGCTCATATTAAAGAAATCCAATTCTCTAGTCACCTCTCTTACAGCGTGGACTGGATCAAGCCAT 1350
 1351 GTTTTTTAATTGTTATTAAAGCAATCAGCATCGGAGACTCACAGCTTTCTACATTGCTCATTGCTCAC 1425
 1426 GTATAGTCAAACCTTATGCTTTCTTGAATTATATCATGCCCTCCATTGTTGAAAAATCTAAAC 1500
 1501 ATTTTTAATGCTCTCCATCAACTAACATTGGGATCATTGAAAGAAAACCTATACAGTTGCTAATACTGCTATAT 1575
 1576 CTCTTTTTCTCTCTTATTGTTTACCCAAAACGTAGCTGGTCTATTCTATCTTGGGGTTCTTAC 1650
 1651 TTCTGACTTCAAGGTATTAGAGCAACTTGAACCAAGATTCTGATCTCAAAAAGCAATAATTATTAAAGTGC 1725
 1726 GATACCAAAGTTCAGAAATAGATCAAACACTCTCCGTGTTTGGCAAATTGTACTGGTAAACCTGAAAC 1800
 40 1801 CGTTGTTTCTAACTCTGAATTAAACCTTTCTTACATTATCTTCTGGACAATGACTTGTCAAATATTG 1875
 1876 TAAAGTGGATTCCAGAATGGTCATAACATATGTCATTGCAAATGGGTGATAAACACAAACTTAGATGAAC 1950
 1951 CGCTAAATTAAAGCACTTCGTGAAATGTTAGATTATTAGCTCTCATAATGTTACAATAAGTGAGGACGACC 2025
 2026 TGAAAGAACTCTCTTTAGTTGTTACTCATGTTGCAAAGGTGAATGAGAAAACAAATAGTATTTCCGACAT 2100
 2101 ATGAAGTACATGTTGTAAGTTGTAACTCATTTGTTACTATTGATGAGCGGTGCTTAAACATCATTGCTTG 2175
 45 2176 ATCACGAATTGTTGTCATTGCTAAAAACCCGACATAATTCTCTGTTAAACATCATTGCTTG 2250

- 43 -

2251 ATGGCTTTTCGGTGGAGCCAGCATTGCTAAACTTAATAAAGAACATGTTAAGTTAACAGAACATTATA 2325
 2326 TGAAAGCATTAGCCAGTACTTCAGATGTTACGTGTTGCTGCAAAAGTTTGCCCCATTTCATTAAGGGAC 2400
 2401 CTAATAATCTTGATATAGTGAATTACAAGGAAAGTAAAGCCTTGATTTAAATACGTGAAAATTTGGCGG 2475
 2476 TGGAAAAATACAGCTATTTAGAAACGCTCATCTTCTGGATCTCCTTATCTAGAGTGGTAGAAGAACAGAAT 2550
 5 2551 TACATTTGTACTATTGGAAGTTATCTCTGTGATAAACAGCGGAATATTTATCAAGGCATTGGCTCAGCG 2625
 2626 CTCTGCAACAAAATGCCCTCGACCGCGTCAATATCCGTTGGCAATTACTTCTCCATATTGGCCAACAGTGTCCG 2700
 2701 TTGCGATTGTCCAAGGTATGGTAAAAAACGAAACATAGCCAGTTATTGCTCAGCTTATGAATATTCCGAGG 2775
 2776 GCGATTTCTTATTCGAACACAGCGTACACTTACCACTTCTTGACTTACTAAAAACAAAGCCTTAATAGTAC 2850
 2851 GTATAGCTGAACCTTCACAAAGTGATGTTGCTACTTTGCTGCCATCCAATATGCATAAAATCCTGCTCGCTAC 2925
 10 2926 TTACTACGGATCATCCTAATTTGGAAGAGAGTGTGATGCTCTTCACTGGCACCTCTGATTTGAAAAAG 3000
 3001 TTGATTTAACGCTTTGTTACGCTCTGATCCTATTCTTACAGTGGAATTGCTTACAGCTTATCAGAACATG 3075
 3076 TTCCCTCATGAAAAAATTGAAAATGCTTAAAGAAAGGTAGCAATGTTGCTCTCAAGTGGTAAATGACGAAGACT 3150
 3151 TGAGCAATAAGGAATTACTTATGATTTTTAATAATCACATTGGGTATCTTAGCAGAATTCTAATATCC 3225
 3226 TTAACGACCTGAAAGGAAAGACTTCATTAATGAAAAGATTAAAGACAATTGCGCAATTGAAAAATGTTATCTT 3300
 15 3301 TATGTGGAGGTGCACTGCTAAACTTGGATTACACAGAATCTTCTAATTTACAAAGTGCTTCTCAAATGAGCT 3375
 3376 TAAGGTTTATGCAATCAAAGCTGGTCACTGATATTAGCAACCAAGGAGCCGAGGATAGTTCAATTGCTG 3450
 3451 GTTAAAGTCTTGTAATTTCCTCTTATTCCCTTATTAGAACACAAGAACAGAGCTAGTAATTCAAATAT 3525
 3526 TTGATTTTATTCCTGACACACACAAGTGCCTACAAGGATTAAGTGGGCTATCCCACCAAGTCTGGATTAG 3600
 3601 CGTGCTTCTAGCCTTAAGGCTAAAGAAATATTCTGTCGCTCAAATGAAGATTTTACTCTGAGCTTCAAAGTA 3675
 20 3676 TAATTAAAGTGTAACTAACGAAAATGAGCCAGTTGTTATTAGGTTACAAAATTAGAACCTTTTCAAG 3750
 3751 CCAAGGTGGACGAGTTACATGACACACTAAATTGGACATATCCAAAGGTTCTGGACCAATTACTAACAGATGCC 3825
 3826 TTTAGATTGTTGTAAAATATGCTTCAACAAATATGCAAATATCATATCTTGCTGCAAAAAATCTGGTGAAT 3900
 3901 TGGGTGCGATAGATCCCAGCCGCCAAGGCTCAACATATTAAAGAAACAGTTGTTCTGATAACTTGAAA 3975
 3976 ACGGAGAAGAAAGTTGAAGTTATTCTAGATTTATGCAATGCAAGTTAATCCAGCTTCTTGTTACTACTG 4050
 25 4051 ATACTAAAGCACAAGGTTTCTGCTATGCTCTGCAAGAGTTCTAAAGCTGGGATTCAAGTCCGAGTGA 4125
 4126 TTAATAAAAAAAAGGGACTAATGTTGCTAACAGAACATTGGATGTTCTGGCTGATTTATCCAAACGTGTGCTTA 4200
 4201 TACCAATTAACTTCAAGTATTTAACACCAATCCCAAATTGACATTGGTACCTGCTACCCATTATTAAGAAA 4275
 4276 ATGTTACTATTCTACACTGGATGCTAGTTGTTCTCTTAAATTGATGGAGTACGCCCTTCGCAAAACGCTGAAA 4350
 4351 AAATATTGGTATTGTTGAAAGTAGTGAAGAACCCAAGGAGTTAACATTCCCTGTTTCTCTCCCTTCTG 4425
 30 4426 TTTAAATGTTATTAAACCGAGTCAGAACACTGGAAGTTAACAGTCAATTGAAAGATTCCAGCTGTTATTAAAC 4500
 4501 AACCGGGACCTGATGGATTAAATTCCGGGGCAACAAAGAACATACCTCATTGAGATGTTATTAAAGATTG 4575
 4576 TGGATTACCTAACAAATGGCTTGCATGCAAGAACAGAGGAAATTGGGATAGACGTTCTGCCATTGCAAGGAAAG 4650
 4651 AGAACCGTTATATGCGGTGGAAAGATGCTACCTCTCGAGAACATCGATCTCAAAGTTGAGTCATTCTTCTC 4725
 4726 GATTTCTTCAAAACATTAGGTATTGTCCTTTAAATTGGGATTTCATGCTCGTGCAATTGTTATTGGGAGC 4800
 35 4801 AACACATACGTAATGCTACAGCTTCAATGCAAGCTTTAGAGTCCGATTATAGAGTTTGCAAGGAAATATGCTG 4875
 4876 GAATTGATGATCCAGATGAAATGCAAGCAGTGTCTTAAATTCCATGATTACTCGTTGATCAACAACTCCTT 4950
 4951 TACATGAAAATTCAAGAACATGGACTCGGCTTGAGTTGCTACGAAATTATTATTCAAAAGGATCCTGAAAATA 5025
 5026 AAAAGGCAGAAATCGGTTGCTAACAGCATGCTGCAATCGGGGATTATGAATCTCTGTTTGAGTTAGATT 5100
 5101 CTTTTATAATCAATGACAACCACCGAGTATCGAAGATGTTAAATTGGGATTGAAGCTTATGGCTATGGCTT 5175
 40 5176 CTATTGATTGTTAAAAAGTGTCTTCAAAAGCAACTTGGAAATCTTCGAAGCTAAATTGGTAGCATATT 5250
 5251 ACCAATACCTACGGAAGGATTCTTGTGAATTGACGGAGCGGCTGCAACCCCTTGACGTTGATGCTGCTACAG 5325
 5326 CAATTGCAAACACAGGCCATTAGCCTATGATTGTTATGATATTATCTAACAGCTGCAACGCAATTAAATGACT 5400
 5401 TTAGTAGGATTGCTGAAACTGACGGAATTGTTCCGACAATCTTGATATTGTTCTCGCCGTCGGCTTACGCAAG 5475
 5476 TAGCTCCGTACGGTAAATTCAAGCAGAACATTCTGTCACACTCACTTAGTTGGCTATGAAAATTGAAAACAGA 5550
 45 5551 AGAAAATGCTGAAATATATCTGAGATTGCAAGAACATCTGAAAAATTGGTCAATTCAAAGAGCCTTCAATG 5625
 5626 CCATCCTCAAAGCAAATGGATTAGATAACCGCTAGCAACAAAGAGCCTTCAATG 5700
 5701 AACATCGTAAAGCTATTCTGAATTGAAATTTCGCTTAATAACACATGTTGATTGGTGTGAGCATGAG 5775

- 44 -

5776 AAAGACCTAAAAATCGTAAAGAACCTTAGGAAATCCACTTAAAGGAAAGTGTCTTGAAACTTACAAAATGGC 5850
 5851 TCGGAAAAGCTGGCAACTGGGATTGAAGGTTGGAGACGTATTATCATAAAGCGTAGAGATTTACTCAGAAT 5925
 5926 GTGAGAATACGCATTATTATCTGGCCATCATCGAGTTTAATGTATGAAGAAGAACAAAAGCTCCAGTTAATG 6000
 6001 AACAGAGCGAACGATTTTAAGTGGTAGGTAGTAACCTCGCATATAAAACGAATTGGTCATCTTGACTATG 6075
 5 6076 GTACAAATCATATATGAAAGTATGCCAAATTGCTCACACTGTGGCTTGATTTGGGCCAGAACCTCGCT 6150
 6151 TATCTAAAGATGACGGCGAAAAGTACTTCTGTGAACACATTATCTCTCGAGAAAAAAATCTTGAACCTTATGA 6225
 6226 ATTCGAATGTTTGTGCCCTTCTATGAAAATTCTCAACTCTTTCTGGTGCATTATCCAAATGATATCCA 6300
 6301 GAGTATGCCATCCAATAATAAAAGTTATAAAAATTGGAACATATAATTGCAAACGTTGTAGCATCTTATCCTG 6375
 6376 GGGAGACGCTATGGCAATTAAATGGCAACAATAAAATCGACTCTCAAAAGCGCTCGCTCGTGGAAAAGCATT 6450
 10 6451 TAAATGTTTACATTCTAGGAAGCTTCTATGTCCTTCAAAGTTGATAAAAAGCACTCAGTCATCTGCAATT 6525
 6526 TCATTACTGAAAAGTTAATCAATTGTGCAATACAAGGATTAACAGTAAATCTGAAAAATGAGCTTAAAGGATC 6600
 6601 ATTTGGCTTCTTTGATGATCCGGTAGATTAGTCATTCTGCTAAATCATTTAGACATTACTTACCAAG 6675
 6676 CTAAAGATGCTAACAGAGCTAGTCATTATCCATTCTCAAACAAACTCAACCTACTCTGTTGAAATTGAGGATGAGG 6750
 6751 TGGATATAATGAACTCTTCAAAACCAAGAAAAGTGTACGTTAGAGGTACGGATGGCAACTTACCCATTCT 6825
 15 6826 TGTGCAAACCCAAAGATGATCTCGTAAGGATGCTAGATTGATGGAATTAAATCTTATTGTTAAATATTGA 6900
 6901 GGAAAGATCAAGAACAGCGAACAGAACAGGAACCTGTGATTAGAAACTATGTTGTTATTCTTAAATGAAGAATGCG 6975
 6976 GATTATCGAATGGGAAATCATACTCGTCCATTAGAGAAAATTGTTAAAAGCTATAGACAGAAAACATTG 7050
 7051 CCATATCATATCAAGAAATCAAAGTTGACTTGCCTAACCCTGGTGTATATTGAAA 7125
 7126 AGAAAATCTACCGAAATTCTCCAGTTTTATGAGTGGTTGTTGAATCTTCCCAGAACAAATAATTGGG 7200
 20 7201 TTACTAGTAGACAAAATTCGCGAACTTAGCAGTAATGTCATAAGTTGGCTACGTTGGGTTGGGAGATC 7275
 7276 GCCATGGCGAAAACATATTGTTGATGAAATTACAGGTGAAGCTATCCATGTCGATTCAACTGTCTTTGATA 7350
 7351 AAGGTCTTACTTTGAAAACCTGAAAAGGTGCCGTTAGATTAACTCATATAATGTTAGATGCAATGGTCCGA 7425
 7426 CAGGTTATGAAGGGGTTTCAAGAACGCTAGCGAAATAACGATGGCTTCTCGCTCAAACCAAGATACTTGA 7500
 7501 TGAGCGTACTAGAGTCCTTACATGATCCTTACTGAGTGGAAATAGAAAAGTCGTCAGCAAGTACCCGA 7575
 25 7576 ATAATGAAGCAAATGAAGTTGGATATAATTGCAAAAAATTCAAGGCTTATGCCAGGGAGACGATACCTT 7650
 7651 TATCTATTGAAGGGCAAATTCAAGAATTGATCAAATCTGCTGTCAACCCAAAAACCTGGTAGAAATGTACATTG 7725
 7726 GTGGGCTGCTTATTCTAGCATTTACTAACAAAATTCAATGAAACAGCTACCCATTATAAAACTTATGATT 7800
 7801 TGAATCGAAGATATTATTATTAAATCCGATGAAAGAATTCTGCTGAGTTGTCATTCTGTAAATTCTT 7875
 7876 CCATTCTAAATCGCTTAAATAGGGCACTGGCTTTTGTGCAATTCTCGTAAAGCAGCTCTG 7950
 30 7951 ATTGAAAAAAAGCTATATCTGTTCTGAGTCATCATCCGAATCAACAAATATAATTGCAAGATCGACCTGCAG 8022

In italics. sequenced by Seaton et al.

In Bold are those bases deleted in Seaton et al. (2499, 22501, 2507, 2509)

Underlined are the two bases either side of a single C insert (5918/5919) in Seaton et al.

(i.e. the incorrect base not shown, but the one residue either side is)

35 Sequence ID No. 4: rad3 protein

1 MSQHAKRKAGSLDLSPRGLDDROAFGQLLKEVLALDKHELGRSNSLPSMTSELVEVLIEVGLLAFKHDDSKSEF 75
 76 ISPMLKEAHLSQLALMILIKRSPVTREIKSSVTLLDWLPRTISLFA DIRFIKLFDSLKEFHKLIVQLISEKS 150
 151 FLWDLYASFMRYWKYYITNVSSIVLQIQTNTATFPYKMPSPNSQPLQISPYPPTHREDKF DLLIINIEACTFFF 225
 226 SAHFFAQCSYLKKSNFPSPPFLTAWTWKPCFFNFVILLKRISIGDSQLFLHLHSRIVOTLCCFSLNFIYHGLPI 300
 40 301 CEKSXHILMSSINLTGSLKKTYTVANTAISSLFFSLFVLPKTVAGLFYPFGVSLLSDFKVLEQLEPDSLKKAI 375
 376 ILFKCRYQSSEIDQTTLRAFGEICTGKLNTLFSNSENLFLHYLSLDNLSNILKVDFOQGHNICTFAKWCIN 450
 451 NNLOEPSNLKHFRMLEDDYSSHNTVISEDDLKNFSLVLCVTHVAKNEKTNISFRTYEVHGCEVCNSFCLLFDERS 525
 526 PFKIPYHELPCALLKNPDIISSSVKQSLLDGFFRWSQHCSNFNKESMLSREFIMKALASTSRCLRVVAAKVLP 600

- 45 -

601 IFIKGPNNLDIVEFHKE SKALIFNTLKILAVENTAILETVLSWISLSRVVEEEELHFVLLLEVVISSVINS GIFYO 675
 676 GIGLSALQOIASTRHISVWQLLSPWPTVSAIVOGMGGKPNIASLFAOLMNISEGDFLIRTOAYTLPFLVLTKN 750
 751 KALIVRIAELSQSDVATLCLTNMHKILASLLTTDHPNLEESVMLLSLATSDFEKVDLTSLLRSOPISITVELLO 825
 826 LYQNDVPHEKIEALRKVAMJVSOVVNDEDLSNKELLYOFFNNHILGJIAEFSNLNDLGKTSINEKIKTIVGI 900
 5 901 EKMLSLCGGAVKGLPOILSNLOSAFONEHLRFYAJKAWSLILATKEPEYSSIAGLSLVJLPPLFPLYEPQAE 975
 976 LVIOIFDFISSLDTKCLOGLKWA IPTSLSACFSLKAKEIFCSLONEDFYSELQSIIKCLTNEPVCYLGLOKL 1050
 1051 ELF FOAKVDELHOTLNLDISNEVLDQOLLRCCLDCV KYASTAQISYLAARNLGE LGAIDPSRAKAQHJIKETVV 1125
 1126 LDNFENGEESLKFILDFMOSQLIPAFLVTTDTKAQFLAYALQEFLKLOGGFKS A VINKKGLTVVTEHMSLPDL 1200
 1201 SKRVLIPFLTSKYHLTPIPKIDIRYPYKENVTIHTW MOLFSLKLMEYAH SQA EKIFGICSKVVKDQE VNIPCF 1275
 10 1276 LLPFLVLANVILTESELEV NKVIEFOLV INOPGPDG LNSVGQ QRTSFYDVFFKIVDYLNKWL RMRKKRNWRRS 1350
 1351 AJARKENRYSVEDATSRESSISKVESFLSRFPSKTLGIVSLNCGFHARALFYEQH IRNATAP YALESDYRVL 1425
 1426 QEIJYAGIOPDEIEAVSLNFHDYSFDQQLL LHENGTHDALSALSCYEIIIQDOPENKKAKIGL NNMSLQSGHYE SL 1500
 1501 VLSLDSFIINDNHEYSKMLNLGIEASWRLSIDS LKKCLSKSNLESFEAKLGSIFYQYLRKDSFAELTERLQPL Y 1575
 1576 VDATAJANTGAHSAYDCYDILSKLHAINDFSRIAETDGIVSDNLDIVLRRRLS QVAPYGF KHOI LSTH L VGYE 1650
 1651 KFENTKKTAEIYLEJARI SRKNGQFORAFNA ILKAM DLDKPLATIEHAQWWHOGQHRKAISELNFSLNNNMFDL 1725
 1726 VDEHEERPKNRKETLGNPLKGKVFLKLTKWLKGAGOLGLKOLETYYHKA E IYSECENTHYYLGHHRVLMYEEEQ 1800
 1801 KLPVNEQSERFLS GELVTRIINEFGRSLYYG TNHYESMPKLLTLWLDGAEELRLS KODGE KYFREHI ISSRKK 1875
 1876 SLELMNSNVCRLSMKI P QYFFLVALS QMISRVCHPNK VYKILEHIIANVASYPGETLWOLMATIKSTSOKRSL 1950
 1951 RGKSILNVLHSRKLMS SKV DIKAL SQSAILITEKLINLCNTRINSKSVKMSLKDHFRLSFDDPV DLVIPA KSFL 2025
 2026 DITLPAKDN RASHYPFPKTQPTLLK FEDEV DIMNSLOPRKVYVRGTDGNLYPFILCKPKDOLRKDARLMEFN NL 2100
 2101 ICKILRKDQEANRRNLCIRTYVVIPLNEECGFIEWVNHTRPFREILLKSYRQKNIPISYQEIKV DLFALRSPNP 2175
 2176 GDIFEKKILPKFPPV FYEWFV E S FPEPNWVTSRQNYCRTLAVMSIVGVVLGLGDRHGENILFDEFTGEAIHVDF 2250
 2251 NCLFDKGLTFEKPEKVPFRLTHNMVDAMGPTGYEGGFRKASEITMRLLSRNQDTLMSVLESFLHDPLV EWRKKS 2325
 2326 SSKYPNNEANEVLDIIRKKFQGFMPGETIPLSIEGQIQELIKSAVNPKNLVEMYIGWAAYF 2386